A Data Sharing Environment for Research & Analysis

Presented by:

Kris L. Hoffman

September, 2012

Agenda

- The Big Challenge
- Reality
- DMDC Data Holdings
- Research and Analysis – Today
- Research and Analysis – Tomorrow
- Person Data Environment
  - What it is
  - Trust Model
  - Data Access Control
- Summary

The Big Challenge

- **Half of the U.S. Defense Budget is spent on people ---- Pay, Benefits, Health Care, Training, Housing, Child Care, etc.**
- **Our people are the most significant weapon in our arsenal --- Yet we make many decisions impacting people based on little analysis e.g.,**
  - **What are the consequences of long or multiple deployments?**
  - **What compensation incentives are most important in shaping retention decisions?**
  - **How can the health care system be strengthened to save money and improve**

Reality

Personnel Research is much harder than it need be:

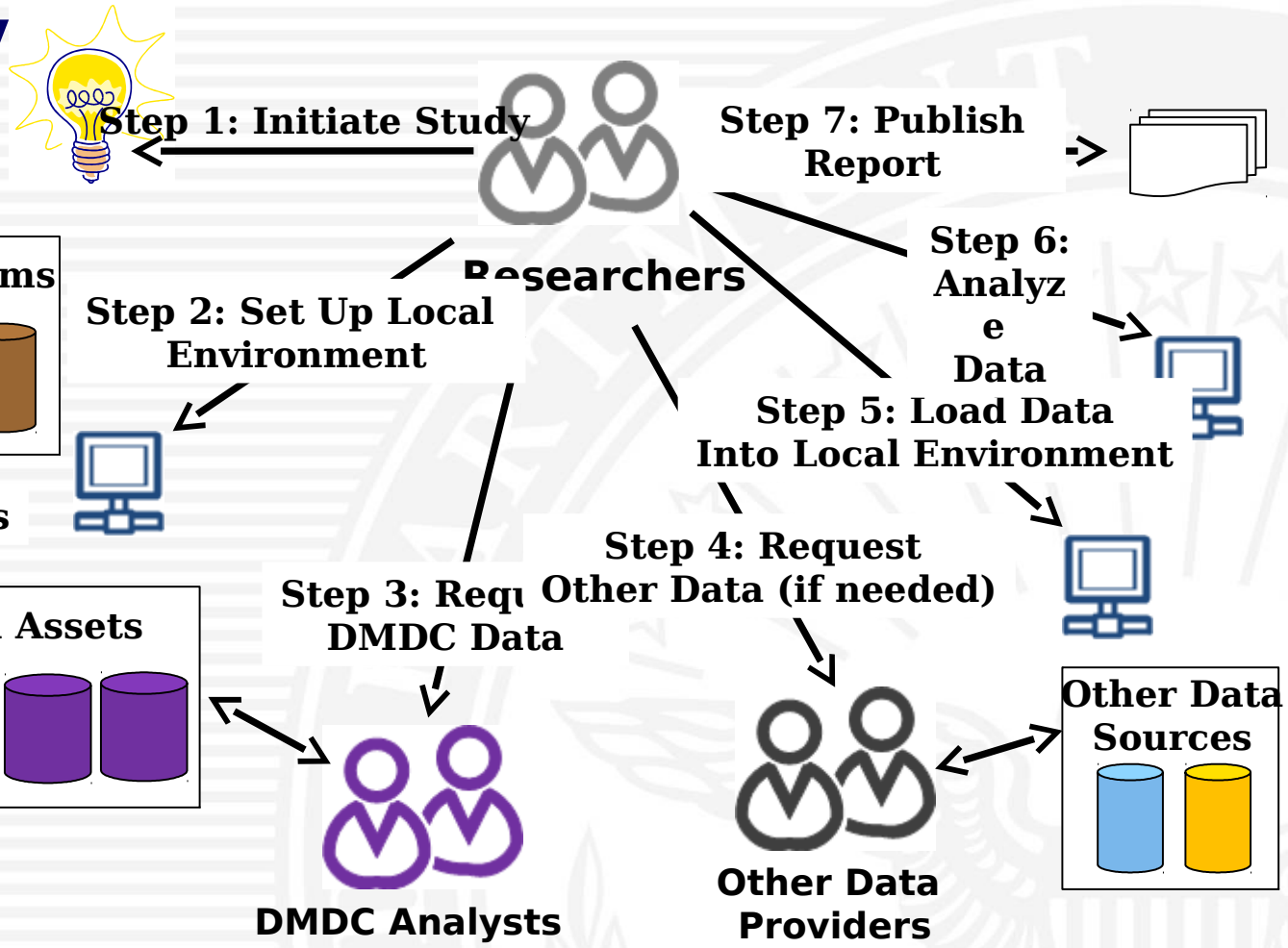
- Each study requires a series of common “data” and “tools” - we lack a central archive to reuse data
- Each study requires data meaning (metadata) knowledge transfer – but we lack tools to share these understandings
- Most studies require cohort/population understanding - but we don’t preserve and share that knowledge
- Most studies require data aggregation knowledge transfer or in some cases development/definition - but we have lost lessons
- Each study requires some level of data quality analysis – but data quality is often a casualty of

DMDC Data Holdings

- Personnel/People Files
- Pay/Financial Files
- Contingency/Operations Files
- Training/Education Files
- Manpower/Force Structure Files
- Medical/Health Files

DMDC Receives/Maintains 91 Data Files ~ About 15K Submission Files Annually

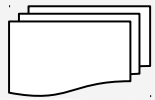
Research and Analysis - Today



Research and Analysis - Tomorr

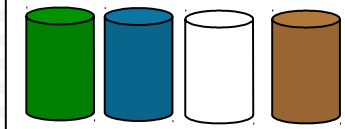
Step 1: Initiate Study

4 X Step 7: Publish Report



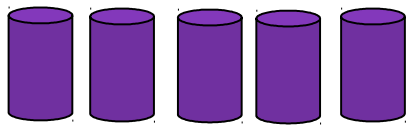
Researchers

Service Systems



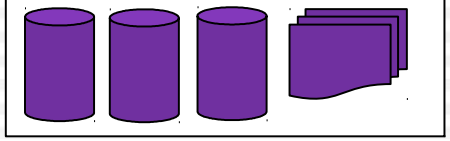
Submissions

DMDC Data Assets



Data Loads

Person Data Environment



Step 2: Set Up Local Environment

Step 3: Request DMDC Data

3 X Step 6: Analyze Data

DMDC Analysts

Step 4: Request Other Data (if needed)

2 X

Other Data Providers

Other Data Sources

Step 5: Load Data Into Local Environment

Step 6: Analyze Data

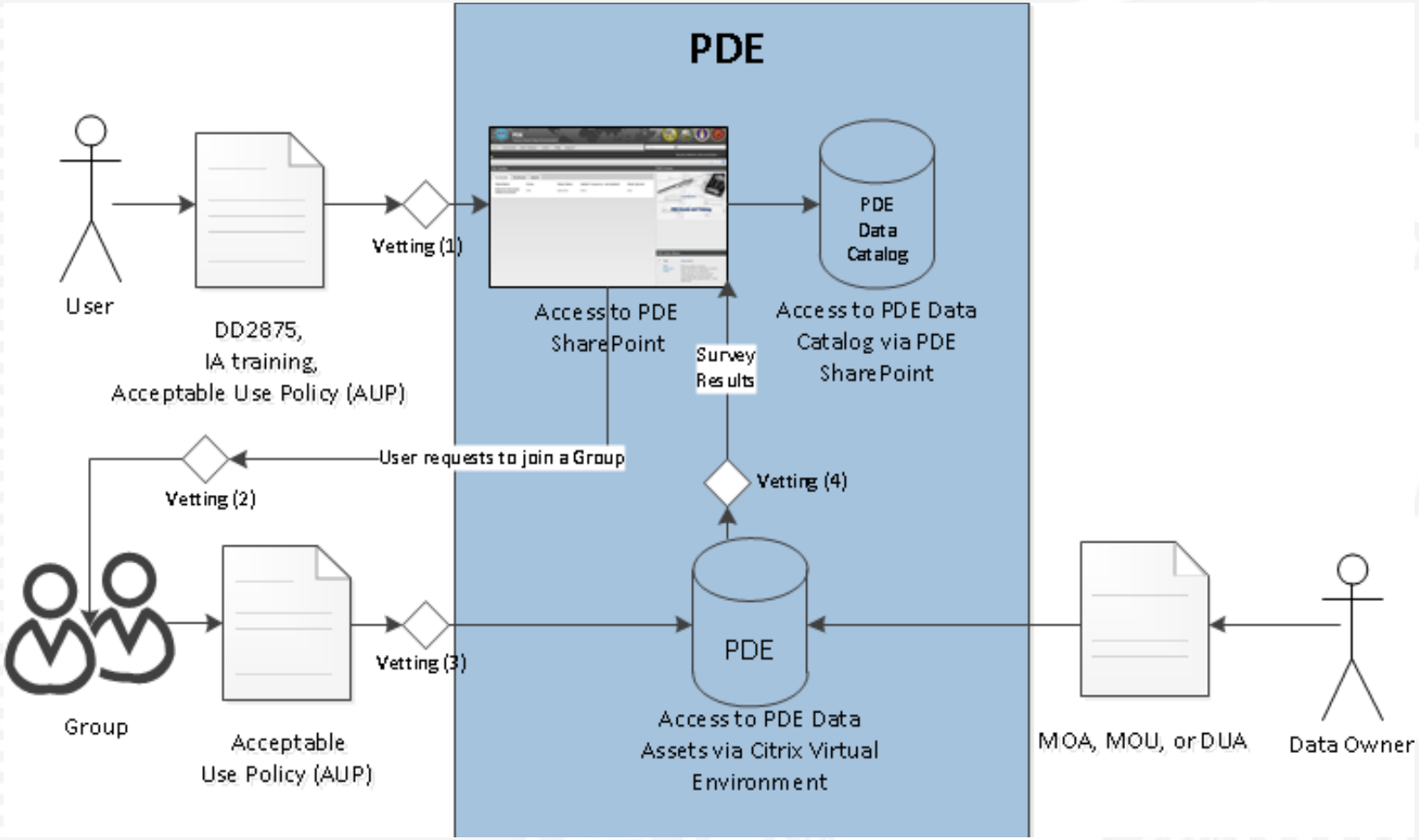
Person Data Environment (PDE)

- Brings the analyst to the data, rather than data to the analyst
- Co-locates various data into a single environment
- Standardizes data elements among disparate person-event databases from across the Department of Defense
- Provides the analyst access to views of the data as required to complete their study
- Complies with the Privacy Act and the Health Information Portability and Accountability Act by replacing person identifiers with

Person Data Environment (PDE)

- Includes a web front end (SharePoint) for collaboration and a virtual desktop environment (Citrix) where the analyst completes their work
- Includes business intelligence, statistical analysis, and web reporting tools through a secure access portal (e.g., SAS, STAT-A, R, SPSS, etc.)

PDE Trust Model



Goal: Establish trust between users, groups, and data owners.

PDE Data Access Control

	Data Catalog			Data Assets			Publish research results to PDE SharePoint
	Public	Restricted	Private	Public	Restricted	Private	
User	Yes	Yes	No	No	No	No	No
Group	Yes	Yes	Only for groups which own the private data asset	Access must be approved by the Group Admin	Access must be approved by the Data Owner	Only for groups which own the private data asset	No
Study	Yes	Yes	Only for studies which own the private data asset	Access must be approved by the Group Admin	Access must be approved by the Data Owner	Only for studies which own the private data asset	Yes*

* The publication of research results to PDE SharePoint may require one or multiple reviews. (e.g. Privacy, Public Health, Human Subjects, Data Owner, PDE Admin)

Summary

- **PDE brings the analyst to the data, rather than data to the analyst**
- **All PDE users can view the Data Catalog and Data Quality information**
- **Members of research groups can view the actual data only if assigned permissions by the Group Admin or Data Owner, depending upon the type of data**
- **Encourages collaboration among researchers, their teams, and the Department of Defense**
- **Goal: Increased collaboration and reduced costs**

Thank You!

POC: Kris L. Hoffman
831-583-2400
Kris.Hoffman@osd.pentagon.mil
My email address changes on September 18, 2012 to kris.l.hoffman.civ@mail.m